Weights in stable marriage problems increase manipulation opportunities

Maria Silvia Pini[1], Francesca Rossi[1], Kristen Brent Venable[1], Toby Walsh[2]
[1] : Department of Pure and Applied Mathematics, University of Padova, Italy.
Email: {mpini,frossi,kvenable}@math.unipd.it.
[2] : NICTA and UNSW, Sydney, Australia. Email: tw@cse.unsw.edu.au

Abstract

The stable marriage problem is a well-known problem of matching men to women so that no man and woman, who are not married to each other, both prefer each other. Such a problem has a wide variety of practical applications, ranging from matching resident doctors to hospitals, to matching students to schools or more generally to any two-sided market. In the classical stable marriage problem, both men and women express a strict preference order over the members of the other sex, in a qualitative way. Here we consider stable marriage problems with weighted preferences: each man (resp., woman) provides a score for each woman (resp., man). In this context, we consider the manipulability properties of the procedures that return stable marriages. While we know that all procedures are manipulable by modifying the preference lists or by truncating them, here we consider if manipulation can occur also by just modifying the weights while preserving the ordering and avoiding truncation. It turns out that, by adding weights, we indeed increase the possibility of manipulating and this cannot be avoided by any reasonable restriction on the weights.

1 Introduction

The stable marriage problem (SM) [3] is a well-known problem of matching the elements of two sets. It is called the *stable marriage* problem since the standard formulation is in terms of men and women, and the matching is interpreted in terms of a set of marriages. Given n men and n women, where each person expresses a strict ordering over the members of the opposite sex, the problem is to match the men to the women so that there are no two people of opposite sex who would both rather be matched with each other than with their current partners. If there are no such people, all the marriages are said to be *stable*. In [1] Gale and Shapley proved that it is always possible to find a matching that makes all marriages stable, and provided a polynomial time algorithm which can be used to find one of two extreme stable marriages, the so-called *male-optimal* or *female-optimal* solutions. The Gale-Shapley algorithm has been used in many real-life scenarios [12], such as in matching hospitals to resident doctors [6], medical students to hospitals, sailors to ships [8], primary school students to secondary schools [13], as well as in market trading.

In the classical stable marriage problem, both men and women express a strict preference order over the members of the other sex in a qualitative way. Here we consider stable marriage problems with weighted preferences (SMWs). In such problems, each man (resp., woman) provides a score for each woman (resp., man). Stable marriage problems with weighted preferences are more general than classical stable marriage problems. Moreover, they are useful in some real-life situations where it is more natural to express scores, that can model notions such as profit or cost, rather than a qualitative preference ordering.

In [10] we have defined new notions of stability for SMWs which depend on the scores given by the agents. In this paper, we study if the stable marriage procedures which return one of these new stable marriages are manipulable. In [11] Roth has shown that, when there are at least three men and three women, every stable marriage procedure is manipulable, i.e., there is a profile in which an agent can mis-report his preferences and obtain a stable marriage which is better than or equal to the one obtained by telling the truth. In this setting, mis-reporting preferences means changing the preference ordering [11] or truncating the preference list [2].

In this paper, we consider a possible additional way of misreporting one's own preferences, which is by just modifying the weights, in a way such that the orderings are preserved and the lists remain complete. We show that it is actually possible to manipulate by just doing this. Thus adding weights makes stable marriage procedures less resistant to manipulation. Moreover, we show that there are no reasonable restrictions on the weights that can prevent such manipulation.

2 Stable marriage problems with weighted preferences

A *stable marriage problem* (SM) [3] of size n is the problem of finding a stable matching between n men and n women. The men and women each have a preference ordering over the members of the other sex. A matching is a one-to-one correspondence between men and women. Given a matching M, a man m, and a woman w, the pair (m, w) is a *blocking* *pair* for M if m prefers w to his partner in M and w prefers m to her partner in M. A matching is said to be *stable* if it does not contain blocking pairs. The sequence of preference orderings of all the men and women is called a *profile*. In the case of the classical stable marriage problem (SM), a profile is a sequence of strict total orders. Given a SM P, there may be many stable matchings for P, and always at least one. The *Gale-Shapley* (GS) *algorithm* [1] is a well-known algorithm that finds a stable matching in polynomial time. Given any procedure f to find a stable matching for an SM problem P, we will denote by $f(P)$ the matching returned by f.

Example 1 *Assume $n = 3$ and let $\{w_1, w_2, w_3\}$ and $\{m_1, m_2, m_3\}$ be respectively the set of women and men. The following sequence of strict total orders defines a profile: $\{m_1 : w_1 > w_2 > w_3$ (i.e., man m_1 prefers woman w_1 to w_2 to w_3); $m_2 : w_2 > w_1 > w_3$; $m_3 : w_3 > w_2 > w_1; w_1 : m_1 > m_2 > m_3$; $w_2 : m_3 > m_1 > m_2$; $w_3 : m_2 > m_1 > m_3$. This profile has two stable matchings: the male-optimal solution which is $\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ and the female-optimal which is $\{(m_1, w_1), (m_2, w_3), (m_3, w_2)\}$.* □

In SMs, each preference ordering is a strict total order over the members of the other sex. More general notions of SMs allow preference orderings to have ties [9; 5; 4]. We will denote with SMT a *stable marriage problem with ties*. A matching M for a SMT is said to be *weakly-stable* if it does not contain blocking pairs. Given a man m and a woman w, the pair (m, w) is a blocking pair for M if m and w are not married to each other in M and each one *strictly prefers* the other to his/her current partner.

A *stable marriage problem with weighted preferences* (SMW) [7] is like a classical SM except that every man/woman gives also a numerical preference value for every member of the other sex, that represents how much he/she prefers such a person. Such preference values are natural numbers and higher preference values denote a more preferred item. The *preference value* for man m (resp., woman w) of woman w (resp., man m) will be denoted by $p(m, w)$ (resp., $p(w, m)$).

Example 2 *Let $\{w_1, w_2\}$ and $\{m_1, m_2\}$ be respectively the set of women and men. An instance of an SMW is the following: $\{m_1 : w_1^{[9]} > w_2^{[1]}$ (i.e., man m_1 prefers woman w_1 to woman w_2, and he prefers w_1 with weight 9 and w_2 with weight 1), $m_2 : w_1^{[3]} > w_2^{[2]}$, $w_1 : m_2^{[2]} > m_1^{[1]}$, $w_2 : m_1^{[3]} > m_2^{[1]}\}$.* □

In [10] we defined two notions of stability for SMWs based on weights. The first one is a simple generalization of the classical notion of stability: a blocking pair is a man and a woman that each prefer to be married to each other more than α with respect to being married to their current partner.

Definition 1 (α-stability) *Let us consider a natural number α with $\alpha \geq 1$. Given a matching M, a man m, and a woman w, the pair (m, w) is an α-blocking pair for M if m prefers w to his partner in M, say w', by at least α (i.e., $p(m, w) - p(m, w') \geq \alpha$), and w prefers m to her partner in M, say m', by at least α (i.e., $p(w, m) - p(w, m') \geq \alpha$). A matching is α-stable if it does not contain α-blocking pairs.*

In Example 2, if $\alpha = 1$, the only α-stable matching is $\{(m_1, w_2), (m_2, w_1)\}$. If instead $\alpha \geq 2$, then all matchings are α-stable.

To find an α-stable matching, it is useful to relate the α-stable matchings of an SMW to the stable matchings of a suitable classical stable marriage problem, so we can use classical stable marriage procedures. Given an SMW P, let us denote with $c(P)$ the classical SM problem obtained from P by considering only the preference orderings induced by the weights of P. If α is equal to 1, then the α-stable matchings of P coincide with the stable matchings of $c(P)$. In general, α-stability gives us more matchings that are stable, since we have a stronger notion of blocking pair. If we denote with $\alpha(P)$ the SMT obtained from an SMW P by setting as indifferent every pair of people whose weight differ for less than α, the α-stable matchings of P coincide with the weakly stable matchings of $\alpha(P)$.

The second notion of stability based on the weights, defined in [10], considers the happiness of a whole pair (a man and a woman) rather than that of each single person in the pair. Thus this notion depends on what we call the strength of a pair, rather than the preferences of each of two members of the pair.

Definition 2 (link-additive stability) *Given a man m and a woman w, the* link-additive strength *of the pair (m, w), denoted by $la(m, w)$, is the value obtained by summing the weight that m gives to w and the weight that w gives to m, i.e., $la(m, w) = p(m, w) + p(w, m)$. Given a matching M, the* link-additive value *of M, denoted by $la(M)$, is the sum of the links of all its pairs, i.e., $\sum_{\{(m,w) \in M\}} la(m, w)$. Given a matching M, a man m, and a woman w, the pair (m, w) is a* link-additive blocking pair *for M if $la(m, w) > la(m', w)$ and $la(m, w) > la(m, w')$, where m' is the partner of w in M and w' is the partner of m in M. A matching is* link-additive stable *if it does not contain any link-additive blocking pair.*

If we consider again Example 2, the pair (m_1, w_1) has link-additive strength equal to 10 (that is, 9+1), while pair (m_2, w_2) has strength 3 (that is, 2+1). The matching $\{(m_1, w_1), (m_2, w_2)\}$ has link-additive value 13 and it is link-additive stable. The other matching is not link-additive stable, since (m_1, w_1) is a link-additive blocking pair.

The reason why we used the terminology *link-additive* is that we compute the strength of a pair, as well as the value of a matching, by using the sum. However, we could use other operators, such as the maximum or the product. If we use the maximum, we will use *link-max* instead of *link-additive*.

Again, we can relate the link-additive (resp., link-max) stable matchings of an SMW to the stable matchings of a suitable classical SM problem. Given an SMW P, let us denote with $Linka(P)$ (resp., $Linkm(P)$) the stable marriage problem with ties obtained from P by taking the preference orderings induced by the link-additive (link-max) strengths of the pairs. Then, a matching is link-additive (resp., link-max) stable iff it is a weakly stable matching of $Linka(P)$. An optimal link-additive (resp. link-max) stable matching is one with maximal link-additive (resp., link-max) value.

3 W-manipulation

We know that, with at least three men and three women, every stable marriage procedure is manipulable [11], i.e., there is a profile where an agent, mis-reporting his preferences, obtains a stable matching which is better than the one obtained by telling the truth. In stable marriage problems, agents can try to manipulate in two ways: by changing the preference ordering [11], or by truncating the preference list [2].

In SMW problems, there is another way of lying: changing the weights. We show this gives the agents an additional power to manipulate even if the manipulator just changes the weights, while preserving the preference ordering and does not truncate the preference list.

A stable marriage procedure f is *w-manipulable* (resp., *strictly w-manipulable*) if there is a pair of profiles p and p' that contain the same preference orderings but differ in the weights of an agent, say w, such that $f(p')$ is better than or equal to (resp., better than) $f(p)$ for w.

4 W-manipulation for α-stability

We first assume that the agents know the value of α.

Theorem 1 *Let α be any natural number > 1. Every procedure which returns an α-stable matching is w-manipulable, and there is at least one procedure which is strictly w-manipulable.*

Proof: Let $\{w_1, w_2\}$ and $\{m_1, m_2\}$ be, respectively, the set of women and men. Consider the following instance of an SMW, say P, $\{m_1 : w_1^{[x+\alpha]} > w_2^{[x]}, m_2 : w_1^{[x+\alpha]} > w_2^{[x]}, w_1 : m_1^{[x+\alpha]} > m_2^{[x+1]}, w_2 : m_1^{[x+\alpha]} > m_2^{[x]}\}$, where x is any value greater than 0. P has two α-stable matchings: $M_1 = \{(m_1, w_1), (m_2, w_2)\}$ and $M_2 = \{(m_1, w_2), (m_2, w_1)\}$. Assume that w_1 mis-reports her preferences as follows: $w_1 : m_1^{[x+\alpha]} > m_2^{[x]}$, i.e., assume that she changes the weight given to m_2 from $x + 1$ to x. Let us denote with P' the resulting problem. P' has a unique α-stable matching, that is M_1, which is the best α-stable matching for w_1 in P. Therefore, it is possible for w_1 to change her weights to get a better or equal result than the one obtained by telling the truth. Also, since P' has a unique α-stable matching, every procedure which returns an α-stable matching returns such a matching. Thus, every procedure is w-manipulable. Moreover, if we take the procedure which returns M_2 in the first profile, this example shows that this procedure is strictly w-manipulable. □

Thus, when using weights, agents can manipulate by just modifying the weights, if they know which α will be used.

Let us now see whether there is any syntactical restriction over the profiles that can prevent this additional form of manipulation. First, we may notice that this manipulation is only related to the fact that some distances between adjacent weights are made larger or smaller. This, depending on the chosen α, may imply that some elements are considered in a tie or ordered in $\alpha(P)$. Thus, a manipulator may introduce a tie that was not in its real preference ordering, or may eliminate a tie from this ordering. Based on this consideration, we can consider restricting our attention to profiles where ties are not allowed. But this would simply mean eliminating the weights, since in this case the α-stable matchings would coincide with the stable matchings of the SM obtained by just forgetting the weights. We can thus consider what happens if we allow at most one tie (that is, a difference less than α) in each preference ordering. Even this strong restriction does not avoid w-manipulation, since the example in the proof of Theorem 1 respects this restriction. A weaker restriction would be to allow at most one tie in the whole profile, but this would mean requiring coordination between the agents or knowing who is the manipulator. Also, again the same example obeys this restriction. Summarizing, if agents know the value of α, there is no way to prevent w-manipulation!

Some hope remains for when α is not known by the agents. Assume that this is the case, but agents know that α is bounded by a certain value, say α_{max}. Unfortunately, again the example in the proof of Theorem 1 (where we replace every α with α_{max}) holds. Thus every procedure is still w-manipulable, and some are also strictly w-manipulable. Also, restricting to at most one tie per agent will not avoid w-manipulation, since again the same example holds.

The most promising case is when agents have no information about α. In this case, we need to define what it means for a procedure to be manipulable: a procedure which returns an α-stable matching is *α-w-manipulable* if it is w-manipulable for all α and it is strictly w-manipulable for at least one α.

Theorem 2 *There is a procedure which returns an α-stable matching which is α-w-manipulable.*

Proof: Let $\{w_1, w_2\}$ and $\{m_1, m_2\}$ be, respectively, the set of women and men. Consider the following instance of an SMW, P, $\{m_1 : w_1^{[3]} > w_2^{[2]}, m_2 : w_2^{[3]} > w_1^{[2]}, w_1 : m_2^{[3]} > m_1^{[2]}, w_2 : m_1^{[3]} > m_2^{[2]}\}$. For every α, P has two α-stable matchings: $M_1 = \{(m_1, w_1), (m_2, w_2)\}$ and $M_2 = \{(m_1, w_2), (m_2, w_1)\}$. When $\alpha = 1$, M_2 is strictly better than M_1 for w_1 in P, while when $\alpha > 1$, M_2 is equally preferred to M_1 for w_1 in P.

Assume that w_1 mis-reports her preferences as follows: $w_1 : m_2^{[3]} > m_1^{[1]}$. Let us denote with P' the problem obtained from P by using this mis-reported preference for w_1. When $\alpha \in \{1, 2\}$, M_2 is strictly better than M_1 for w_1 in P', while when $\alpha > 2$, M_2 is equally preferred to M_1 for w_1 in P'.

Let us consider a procedure, that we call mGS, which works as the Gale-Shapley algorithm over all the profiles except on P and P', where it works as follows: if a matching is strictly better than another matching in terms of α for w_1, then it returns the best one, while if a matching is equally preferred to another matching in terms of α for w_1, then it returns the worst one for w_1 w.r.t. the strict preference ordering induced by the weights. Therefore, when $\alpha = 1$, mGS returns M_2 in both P and P', when $\alpha = 2$ mGS returns M_1 in P and M_2 in P', while when $\alpha > 2$ mGS returns M_1 in both P and P'. Therefore, if w_1 lies, for every α, he obtains a partner that is better than or equal to the one obtained by telling the truth, and there is a value α (i.e., α=2) where he obtains a partner that is better than the one obtained by telling the truth. Therefore, the mGS procedure is α-w-manipulable. □

As in the case when α is known, we may consider restricting to profiles with at most one tie per agent. However, the example in the above proof satisfies this restriction, so it shows that α-w-manipulability is possible also with such a severe restriction.

Summarizing, in the context of α-stability, no matter whether we have information about α or not, it is possible to have w-manipulability, even if we severely restrict the profiles.

5 W-manipulation for link-additive stability

We next show that every procedure for link-additive stability is strictly w-manipulable.

Theorem 3 *Every procedure that returns a link-additive stable matching is strictly w-manipulable.*

Proof: Let $\{w_1, w_2\}$ and $\{m_1, m_2\}$ be, respectively, the set of women and men. Consider the following instance of an SMW, say P: $\{m_1 : w_2^{[6]} > w_1^{[4]}, m_2 : w_2^{[5]} > w_1^{[4]}, w_1 : m_1^{[4]} > m_2^{[3]}, w_2 : m_1^{[3]} > m_2^{[2]}\}$. P has a unique link-additive stable matching, which is $M_1 = \{(m_1, w_2), (m_2, w_1)\}$. Assume that w_1 mis-reports her preferences as follows: $w_1 : m_1^{[5000]} > m_2^{[2]}$. Then, in the new problem, that we call P', there is only one stable matching, which is $M_2 = \{(m_1, w_1), (m_2, w_2)\}$, and M_2 is better than M_1 for w_1 in P. Since there is only one stable matching in both P and P', every procedure which returns a link-additive stable matching will return M_2 in P and M_1 in P', and thus it is strictly w-manipulable. □

The example in the proof of the above theorem shows a very intuitive and dangerous manipulation scheme: the manipulator sets a very high weight (higher than twice the highest of the other weights in the profile) for its top choice. In this way, it will surely be matched to its top choice, no matter the procedure used or the preferences of the other agents over the alternatives that are not their top choices.

This form of manipulation can be avoided by forcing the same weight for all top choices of all agents. This restriction however does not prevent all forms of w-manipulation.

Theorem 4 *If we restrict to profiles with the same weight for all top choices, every procedure that returns a link-additive stable matching is w-manipulable, and there is at least one procedure which is strictly w-manipulable.*

Proof: Let $\{w_1, w_2, w_3\}$ and $\{m_1, m_2, m_3\}$ be, respectively, the set of women and men. Consider the following instance of an SMW, P, $\{m_1 : w_3^{[7]} > w_2^{[6]} > w_1^{[5]}, m_2 : w_3^{[7]} > w_2^{[6]} > w_1^{[5]}, m_3 : w_3^{[7]} > w_2^{[6]} > w_1^{[5]}, w_1 : m_3^{[7]} > m_1^{[5]} > m_2^{[4]}, w_2 : m_3^{[7]} > m_1^{[5]} > m_2^{[4]}, w_3 : m_3^{[7]} > m_1^{[6]} > m_2^{[5]}\}$. P has an unique link-additive stable matching, which is $M_1 = \{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$. Assume that w_1 mis-reports her preferences as follows: $w_1 : m_3^{[7]} > m_1^{[6]} > m_2^{[4]}$. Then, in the new problem, that we call P', there are only two link-additive stable matchings, i.e., M_1 and $M_2 = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$, where M_2 is better than M_1 for w_1. Thus every procedure is w-manipulable. If we consider the procedure that returns matching M_2 in P', this pair of profiles shows that this procedure is strictly w-manipulable. □

Notice that, if we consider profiles where all top choices have the same weight and all differences (of weights of adjacent items in the preference lists) are exactly 1, then weights are fixed and are thus irrelevant. Also, obviously w-manipulation cannot occur, since agents cannot modify the weights. We may wonder whether, by restricting to profiles which are close to this extreme case, we may avoid w-manipulation. Unfortunately, this is not so. In fact, we can consider just profiles with the same weight for all top choices and where at most one difference is 2, while all the others are 1, for every agent. This holds for the example in the proof of Theorem 4. This shows that even this strong restriction is not enough to avoid w-manipulation.

If we restrict our attention to procedures that return optimal link-additive or link-max stable matchings, we can still prove that all such procedures are strictly w-manipulable, and they are w-manipulable when all top choices have the same weight. In fact, the same examples in the proofs of Theorem 3 and 4 still hold.

6 Conclusions and future work

We have investigated the manipulation properties of stable marriage problems with weighted preferences, and considered two different notions of stability. We have shown that, in both cases, adding weights to classical stable marriage problems increases the possibility of manipulating the resulting matching, since agents can manipulate even by just modifying the weights, without changing or truncating the preference lists. We have also shown that reasonable restrictions over the weights do not avoid such additional forms of manipulation. However, in the case of link-additive stability, forcing all top choices to have the same weight for all agents prevents an extreme form of w-manipulation, which would allow the manipulator to dictate its own partner in every link-additive stable matching.

We plan to investigate the computational complexity of w-manipulation. We also plan to use scoring-based voting rules to choose among the stable matchings, and to adapt existing results about manipulation complexity for such voting rules to weighted stable marriage problems.

References

[1] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *Amer. Math. Monthly*, 69:9–14, 1962.

[2] D. Gale and M. Sotomayor. Semiring-based constraint solving and optimization. *American Mathematical Monthly*, 92:261–268, 1985.

[3] D. Gusfield and R. W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, Boston, Mass., 1989.

[4] M. Halldorsson, R. W. Irving, K. Iwama, D. Manlove, S. Miyazaki, Y. Morita, and S. Scott. Approximability results for stable marriage problems with ties. *Theor. Comput. Sci*, 306(1-3):431–447, 2003.

[5] R. W. Irving. Stable marriage and indifference. *Discrete Applied Mathematics*, 48:261–272, 1994.

[6] R. W. Irving. Matching medical students to pairs of hospitals: a new variation on an old theme. In *Proc. ESA'98*, volume 1461 of *LNCS*, pages 381–392. Springer-Verlag, 1998.

[7] R. W. Irving, P. Leather, and D. Gusfield. An efficient algorithm for the "optimal" stable marriage. *J. ACM*, 34(3):532–543, 1987.

[8] J. Liebowitz and J. Simien. Computational efficiencies for multi-agents: a look at a multi-agent system for sailor assignment. *Electonic government: an International Journal*, 2(4):384–402, 2005.

[9] D. Manlove. The structure of stable marriage with indifference. *Discrete Applied Mathematics*, 122(1-3):167–181, 2002.

[10] M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Stability in matching problems with weighted preferences. In *Proc. ICAART'11*. SciTePress, 2011.

[11] A. E. Roth. The economics of matching: Stability and incentives. *Mathematics of Operations Research*, 7:617–628, 1982.

[12] A. E. Roth. Deferred acceptance algorithms: History, theory, practice, and open questions. *International Journal of Game Theory, Special Issue in Honor of David Gale on his 85th birthday*, 36:537–569, 2008.

[13] Chung-Piaw Teo, Jay Sethuraman, and Wee-Peng Tan. Gale-shapley stable marriage problem revisited: Strategic issues and applications. *Manage. Sci.*, 47(9):1252–1267, 2001.